

# Implementasi Metode *Naive Bayes* untuk Pemilihan Jenis Tanaman Reboisasi berdasarkan Kondisi Lahan (Studi Kasus Reboisasi Kawasan Hutan Sulawesi Selatan)

Suwatri Jura<sup>1</sup>, Abdul Rochman<sup>2</sup>, Billy Eden William Asrul<sup>3</sup> dan Sitti Zuhriyah<sup>4</sup>

<sup>1,2,3</sup> Teknik Informatika, STMIK Handayani, <sup>4</sup> Sistem Komputer, STMIK Handayani

<sup>1</sup>[Suwatrijura@handayani.ac.id](mailto:Suwatrijura@handayani.ac.id), <sup>2</sup>[arochmanjasadi@gmail.com](mailto:arochmanjasadi@gmail.com), <sup>3</sup>[billy@handayani.ac.id](mailto:billy@handayani.ac.id),

<sup>4</sup>[Zuhriyah@handayani.ac.id](mailto:Zuhriyah@handayani.ac.id)

Copresponder Author : [billy@handayani.ac.id](mailto:billy@handayani.ac.id)

**Abstract** — Plants for reforestation are planted and maintained with the aim of repairing damaged forests or non-forests in an area of land, where the land can be used by humans to carry out activities to maintain forest ecosystems. To determine the type of reforestation plants that are suitable for land conditions, there are several difficulties encountered. This study aims to implement the Naive Bayes method to adapt plant benefits to damaged land conditions, which include topography, climate, and soil type. The process in this study was carried out from the data cleaning, data integration, and classification stages. The test was carried out in 2 stages, the test was carried out to determine the accuracy of the Naive Bayes Model for the Data and Application Output Results. The data was tested with 10-fold validation. Based on the results of the scenarios for the two testing stages carried out in this study, the first scenario obtained an accuracy rate of 96% and the accuracy of the second stage of testing was 83.53%, the average accuracy of the two testing stages obtained results of 89.76%.

**Keyword** — Jenis Tanaman, Lahan, *Naive Bayes*, Reboisasi.

**Abstrak** — Tanaman untuk reboisasi ditanam dan dipelihara dengan tujuan memperbaiki hutan atau non hutan yang mengalami kerusakan pada suatu area lahan, dimana lahan dapat dimanfaatkan oleh manusia untuk melakukan kegiatan menjaga ekosistem hutan. Untuk menentukan jenis tanaman reboisasi yang sesuai dengan kondisi lahan, terdapat beberapa kesulitan yang dihadapi. Penelitian ini bertujuan mengimplementasikan metode *Naive Bayes* untuk menyesuaikan manfaat tanaman dengan kondisi lahan yang rusak, yang meliputi topografi, iklim, dan jenis tanah. Proses pada penelitian ini dilakukan dari tahap *data cleaning*, *data integration*, dan *Classification*. Pengujian dilakukan dengan 2 Tahap, pengujian dilakukan untuk menentukan akurasi dari Model *Naive Bayes* terhadap data dan hasil luaran Aplikasi. Data diuji dengan *10-fold validation*. Berdasarkan Hasil skenario kedua tahap pengujian yang dilakukan pada penelitian ini, diperoleh tingkat akurasi skenario pertama sebesar 96% dan akurasi pengujian tahap kedua sebesar 83.53%, rata-rata akurasi dari kedua tahap pengujian didapatkan hasil sebesar 89,76%.

**Kata kunci** — Jenis Tanaman, Lahan, *Naive Bayes*, Reboisasi.

## I. PENDAHULUAN

Reboisasi merupakan suatu usaha yang dilakukan masyarakat untuk memperbaiki hutan atau area non-hutan yang telah mengalami kerusakan yaitu tandus atau gundul, dimana hutan mengalami kerusakan karena beberapa alasan,

seperti kebakaran hutan, penebangan liar, longsor, dan erosi tanah, dengan cara penanaman kembali tanaman atau pohon pada kawasan tersebut [1]. Tanaman untuk reboisasi ditanam dan dipelihara dengan tujuan memperbaiki hutan atau non hutan yang mengalami kerusakan pada suatu area lahan, dimana lahan dapat dimanfaatkan oleh manusia untuk melakukan kegiatan menjaga ekosistem hutan. Tanaman reboisasi biasanya digunakan untuk menjaga keseimbangan ekosistem tumbuhan bagi lahan yang dimanfaatkan.[2].

Berdasarkan data dari Badan Lingkungan Hidup daerah provinsi Sulawesi Selatan pada tahun 2015 mengalami kerusakan hutan seluas 138.654,00 Ha. Penyebab kerusakan hutan terbesar adalah kegiatan lainnya seluas 131.816 Ha (95,07%), perambahan hutan seluas 6.653,50 ha (4,80 %), kebakaran hutan seluas 169,50 ha (0,12 %), dan akibat penebangan liar seluas 15,00 Ha (0,01 %). Upaya pengelolaan lingkungan sebagai penanggulangan pada kondisi lingkungan yang mengalami kerusakan tersebut pun dilakukan salah satunya dengan rehabilitasi lingkungan yakni dengan melakukan penanaman tanaman hijau untuk kebutuhan reboisasi [3]. Untuk menentukan jenis tanaman reboisasi yang sesuai dengan kondisi lahan, terdapat beberapa kesulitan yang dihadapi. Selain karena banyaknya jenis tanaman yang bisa dipilih dan manfaat yang dimiliki oleh tanaman tersebut, juga karena kondisi alam yang menentukan tanaman tersebut cocok atau tidak untuk tumbuh dilahan tersebut [4][5].

Penelitian ini bertujuan untuk menguji apakah metode *Naive Bayes* dapat digunakan sebagai sistem klasifikasi kesesuaian jenis lahan terhadap jenis tanaman reboisasi [6],[7],[8]. Penelitian ini juga bertujuan mengimplementasikan metode *Naive Bayes* untuk menyesuaikan manfaat tanaman dengan kondisi lahan yang rusak, yang meliputi topografi, iklim, dan jenis tanah pada daerah sulawesi selatan [9][10]. Metode *Naive Bayes* dipilih sebagai metode penghitungan dikarenakan metode ini menghitung probabilitas menggunakan data masa lalu dan hanya membutuhkan data latih yang kecil yang banyak digunakan pada kasus pemilihan lahan untuk tanaman . Sesuai dengan data penelitian yang digunakan yaitu data

klasifikasi kesesuaian jenis tanah terhadap jenis tanaman, maka data pengelolaan reboisasi beberapa tahun sebelumnya digunakan sebagai nilai perbandingan untuk mengetahui tingkat kesesuaian jenis tanah terhadap jenis tanaman. Proses pada penelitian ini dilakukan dari tahap *data cleaning*, *data integration*, dan *Classification* data tanaman reboisasi, mengelompokkan suatu class berdasarkan atribut topografi, dan jenis lahan [], sampai dengan menentukan kesesuaian lahan dengan jenis tanaman menggunakan *Naive Bayes*.

II. TINJAUAN PUSTAKA

A. Klasifikasi/Classification

Klasifikasi merupakan proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui [11],[11],[12]. Terdapat beberapa penelitian *Classification* jenis lahan terhadap tanaman sebelumnya, dengan judul Penerapan Model Klasifikasi Clustering Naive Bayes Kesesuaian Lahan Tanaman [4] penelitian tersebut menyimpulkan berdasarkan data yang telah dimasukkan dalam model naive bayes adalah *likelihood* “ya” 0,00036864, *Likelihood* “Tidak” 0,00004608 dan Probalitas “Ya” 0,88888889 dan Probalitas “Tidak” 0,11111111 sehingga kesimpulannya adalah wilayah yang akan di tanam tanaman ubi kayu sesuai karena nilai probalitas “ya” lebih besar dibandingkan nilai probalitas “tidak”. Penelitian selanjutnya dengan judul Pemetaan dan Klasifikasi Kesesuaian Jenis Tanah Terhadap Tanaman Menggunakan Metode *Naive Bayes* di Desa Cukilan [1] Terlihat dari hasil produksi yang berjalan positif dengan hasil probabilitas kesesuaian yang cukup tinggi yaitu jagung 88% dengan rata-rata produksi 4.272,12 ton, kemudian padi 79,52% dengan rata-rata produksi 35.373,94 ton. Dari penelitian sebelumnya maka peneliti mengimplementasikan metode klasifikasi untuk menentukan jenis tanaman reboisasi yang sesuai dengan lahan yang ada pada hutan disulawesi selatan.

B. Naive Bayes Classifier

Klasifikasi *Naive Bayes* merupakan sebuah metode klasifikasi berdasarkan teorema Bayes yang menggunakan metode probabilitas dan statistik yang dikemukakan oleh seorang ilmuwan Inggris bernama Thomas Bayes, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Probabilitas adalah kemungkinan terjadinya suatu peristiwa antara 0 s/d 1 [1],[4],[14]. Klasifikasi Gaussian *Naive Bayes*

dapat digunakan untuk memproses atribut numerik pada layanan jaringan komputer [1], [4], [14].

Persamaan dari teorema Bayes adalah :

$$P(H | X) = \frac{P(X|H).P(H)}{P(X)} \dots\dots\dots (2.1)$$

Keterangan :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data merupakan suatu class spesifik
- P(H | X): Probabilitas hipotesis H berdasar kondisi X (posteriori probability)
- P(H) : Probabilitas hipotesis H (prior probability)
- P(X | H): Probabilitas berdasarkan kondisi X pada hipotesis H
- P(X) : Probabilitas X

Untuk menjelaskan teorema *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema bayes di atas disesuaikan sebagai berikut :

$$P(C | F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n | C)}{P(F_1 \dots F_n)} \dots\dots\dots (2.2)$$

Dimana Variabel C merepresentasikan kelas, sementara variabel F1...Fn merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik karakteristik sampel secara global (disebut juga *evidence*) []. Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut :

$$Posterior = \frac{Prior \times likelihood}{evidence} \dots\dots\dots (2.3)$$

Nilai Evidence selalu tetap untuk setiap kelas pada satu sampel. Nilai dari posterior tersebut nantinya akan dibandingkan dengan nilai nilai posterior kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan[1], [4], [6], [14]. Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan menjabarkan menggunakan aturan perkalian sebagai berikut :

$$P(C | F_1 \dots F_n) = P(C)P(F_1 \dots F_n | C)$$

$$\begin{aligned}
 &= P(C)P(F_1 | C)P(F_2, \dots, F_n | C, F_1) \\
 &= P(C)P(F_1 | C)P(F_2 | C, F_1)P(F_3, \dots, F_n | C, F_1, F_2) \\
 &= P(C)P(F_1 | C)P(F_2 | C, F_1)P(F_3 | C, F_1, F_2)P(F_4, \dots, F_n | C, F_1, F_2, F_3) \\
 &= P(C)P(F_1 | C)P(F_2 | C, F_1)P(F_3 | C, F_1, F_2) \dots P(F_n | C, F_1, F_2, F_3 \dots F_{n-1})
 \end{aligned}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan [1], [4], [6], [14]. Disinilah digunakan asumsi independensi yang sangat tinggi (naif), bahwa masing masing petunjuk saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(P_i | F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \dots \dots \dots (2.4)$$

Untuk  $i \neq j$ , sehingga

$$P(F_i | C, F_j) = P(F | C) \dots \dots \dots (2.5)$$

Dari persamaan diatas dapat disimpulkan bahwa asumsi independensi naif tersebut membuat syarat peluang menjadi sederhana, sehingga perhitungan menjadi mungkin untuk dilakukan. Selanjutnya, penjabaran  $P(C | F_1 \dots F_n)$  dapat disederhanakan menjadi:

$$\begin{aligned}
 P(C | F_1 \dots F_n) &= P(C)P(F_1 | C)P(F_2 | C)P(F_3 | C) \dots \\
 &= P(C) \prod_{i=1}^n P(F_i | C) \dots \dots \dots (2.6)
 \end{aligned}$$

C. Tahap Algoritma Naive Bayes

Dasar algoritma pada Naive Bayes, adalah sebagai berikut :

1) Menghitung jumlah kelas / label

Menentukan kelas/label berdasarkan data training lalu menghitung jumlah total kelas/label [6]. Jika atribut/kasus setiap kelas memiliki banyak klasifikasi, maka dilakukan perhitungan untuk mengetahui nilai pada setiap atribut terlebih dahulu berdasarkan perhitungan nilai dari setiap klasifikasi yang ada [1], [8].

2) Menghitung jumlah kasus per kelas

Jumlah dari masing-masing kelas/label dibagi dengan jumlah total kelas/label yang telah dihitung untuk mengetahui probabilitas dari masing-masing kelas/label [8].

3) Kalikan semua variabel kelas

Setiap variabel kelas/label berdasarkan data training yang telah ditentukan nilainya dikalikan untuk mendapatkan hasil nilai dari setiap variabel kelas yang ada [1], [8].

4) Bandingan hasil per kelas

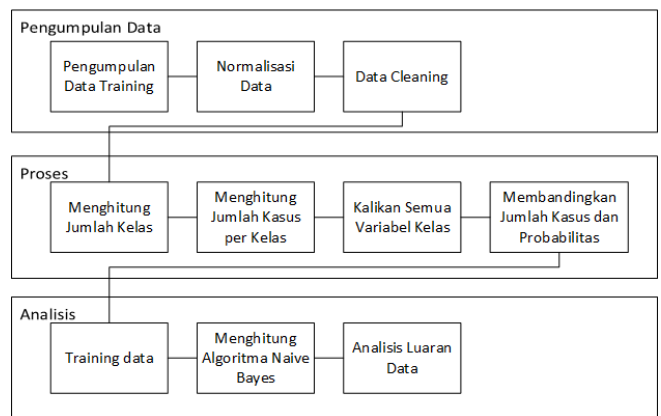
Hasil dari setiap variabel kelas yang telah dihitung kemudian dibandingkan dengan variabel yang memiliki nilai tertinggi/maksimum dipakai sebagai hasil akhir dari kelas yang diinginkan [1], [4], [6], [14].

D. Data Cleaning

Pada tahapan Data Cleaning pengumpulan data tentang hasil konsultasi lahan pada dinas kehutanan dari tahun sebelumnya dilakukan untuk mendapatkan *data training*, yang kemudian pada tahapan ini dilakukan Cleaning data terkait data yang redundansi, data yang tidak memiliki class, data yang tidak lengkap terkait parameter yang jadikan inputan pada penelitian ini, dimana parameter yang digunakan diantaranya: jenis tanaman, manfaat, ketinggian lahan, kemiringan lahan, Curah hujan lokasi, temperatur, tekstur tanah, ph tanah, dan kondisi drainase. Dari data tersebut kemudian didapatkan jumlah *class* jenis tanaman yang akan dijadikan acuan dalam menentukan klasifikasi jenis tanaman untuk lahan baru yang dinilai [8], [12]–[14].

III. METODE PENELITIAN

Penelitian melibatkan beberapa tahapan dalam melakukan Pengumpulan data dan melakukan Perancangan Algoritma yang mengimplementasikan Metode *Naive Bayes*. Tahapan tersebut diuraikan pada diagram alir metode penelitian berikut.



Gambar 1. Diagram Alir Penelitian

A.. Tahap Pengumpulan Data

Wawancara perlu dilakukan mengetahui kebutuhan apa saja yang dibutuhkan dalam menyelesaikan masalah yang telah dipaparkan pada rumusan masalah untuk membuktikan bahwa sistem yang akan dibangun dapat menentukan jenis tanaman yang sesuai berdasarkan kondisi lahan [1], [8]. Mengumpulkan jenis Tanaman reboisasi yang akan digunakan sebagai penentuan *Class* [10], seperti tanaman yang sering dipakai yaitu tanaman Angsana, Mahoni, Pinus, Bambu, Jati, Terambesi, Jabon, Johar, Sengon, Cempaka,

dan Bakau. Dan mengumpulkan data training yang berasal dari data hasil konsultasi reboisasi pada 2 tahun terakhir sebanyak 11 Data yang menjadi Class.

**B. Normalisasi Data**

Data Yang telah dikumpulkan kemudian dilakukan normalisasi untuk mengetahui mana saja data yang dapat digunakan untuk menjadi data training [3]. Dari 367 data dilakukan normalisasi data dengan mempertimbangkan mana saja data yang lengkap sesuai dengan Atribut yang dijadikan inputan [11], setelah itu dilakukan pembersihan data terkait data redundansi [15], dan pada tahap akhir ditemukan 249 data yang akan digunakan untuk menjadi data training. Dari Kumpulan Data dan variabel digunakan jenis tanaman, manfaat tanaman, ketinggian lahan, curah hujan lokasi, temperatur (Suhu), Tekstur Tanah, PH tanah, Dan Drainase untuk atribut parameter perhitungan Naive Bayes.

**IV. HASIL DAN PEMBAHASAN**

**A. Hitung Jumlah Label**

Tahap Awal Dilakukan perhitungan Jumlah Label Data yang didapatkan dari hasil Normalisasi kumpulan Data Training sebanyak 11 Data. Dapat dilihat pada tabel 1.

Tabel 1. Data class dan label

Jenis Tanaman	Manfaat	Ketinggian (mdpl)	Kemiringan	Curah Hujan (mm/th)	Temperatur	Tekstur	pH	Drainase
Angsana	Klimatologis	<1000	0-8%	1000-2000	19-31	Sedang-Berat	Asam Netral	Baik
Mahoni	Klimatologis	50-1400	>45%	1600-4000	18-34	Sedang-Berat	Basa Netral	Baik
Pinus	Klimatologis	300-3000	15-25%	2000-3000	16-30	Ringan	Asam Netral	Baik
Bambu	Hidrologis	<1000	10-45%	1000-3000	8-36	Ringan-Berat	Basa Netral	Baik
Jati	Hidrologis	<900	25-45%	1250-3000	18-32	Sedang-Berat	Basa Netral	Baik
Trembesi	Hidrologis	<300	0-10%	600-3000	20-38	Ringan-Berat	Asam Netral	Baik
Jabon	Ekologis	<1000	8-15%	1300-4000	19-33	Ringan-Berat	Asam Netral	Baik
Johar	Ekologis	<600	15-25%	650-1500	13-35	Ringan-Berat	Asam Netral	Baik
Sengon	Ekologis	<1600	16-30%	2000-4000	20-34	Ringan-Berat	Asam Netral	Baik
Cempaka	Orologis	1500-2500	25-45%	1400-2600	20-25	Sedang-Berat	Asam Netral	Baik

Dari data label diatas, kemudian dilakukan perhitungan data dengan dan didapatkan hasil Berikut adalah kasus data tanaman berdasarkan karakteristik lahan Menghitung jumlah kelas /label

Diketahui class hasil berdasarkan data training di atas memiliki beberapa klasifikasi, yaitu :

- Class1 (C1) = Angsana = 25
- Class2 (C2) = Mahoni = 27
- Class3 (C3) = Pinus = 24
- Class4 (C4) = Bambu = 30
- Class5 (C5) = Jati = 26
- Class6 (C6) = Trembesi = 22
- Class7 (C7) = Jabon = 19
- Class8 (C8) = Johar = 25
- Class9 (C9) = Sengon = 28
- Class10 (C10) = Cempaka = 16
- Class11 (C11) = Bakau = 7

Jadi, jumlah kelas (C) sebanyak 11 Class dari 25 + 27 + 24 + 30 + 26 + 22 + 19 + 25 + 28 + 16 + 7 = 249 data

Karena setiap atribut memiliki beberapa kriteria, maka dilakukan perhitungan untuk mengetahui nilai pada setiap kriteria aribut sebagai berikut :

Kriteria X1= Atribut ketinggian (mdpl)

Diketahui Atribut Ketinggian berdasarkan data training di atas yaitu 225 Mdpl, maka :

Kriteria Inputan X1(225)

Angsana= 0; Mahoni= 0; Pinus= 0; Bambu= 1; Jati= 0; Trembesi= 0; Jabon= 0; Johar= 0; Sengon= 0; Cempaka= 0; Bakau= 0; **Total X1 (225)= 1**

Kriteria Inputan X2 = Curah Hujan

Diketahui Atribut Curah Hujan berdasarkan data training di yaitu 1070 mm/thn, maka :

X2 (1070)

Angsana= 0; Mahoni= 0; Pinus= 0; Bambu= 1; Jati= 0; Trembesi= 0; Jabon= 0; Johar= 0; Sengon= 0; Cempaka= 0; Bakau= 0; **Total X1 (1070)= 1**

Kriteria Inputan X3 = Temperatur

Diketahui Atribut Temperatur berdasarkan data training di yaitu 37°C, maka :

X3 (37°)

Angsana= 0; Mahoni= 0; Pinus= 0; Bambu= 1; Jati= 0; Trembesi= 1; Jabon= 0; Johar= 0; Sengon= 0; Cempaka= 0; Bakau= 0; **Total X1 (37°)= 2**

Kriteria Inputan X4 = Kemiringan

Diketahui Atribut Kemiringan berdasarkan data training di yaitu 41%, maka :

X3 (41%)

Angsana= 0; Mahoni= 0; Pinus= 0; Bambu= 1; Jati= 2; Trembesi= ; Jabon= 0; Johar= 0; Sengon= 0; Cempaka= 0; Bakau= 0; **Total X1 (41%)= 3**

Kriteria Inputan X5 = Tekstur

Diketahui Atribut Tekstur berdasarkan data training di yaitu Ringan-Berat, maka :

X5(Ringan-Berat)

Angsana= 4; Mahoni= 6; Pinus= 4; Bambu= 7; Jati= 4; Trembesi=5; Jabon= 4; Johar= 12; Sengon= 10; Cempaka= 2; Bakau= 5; **Total X1 (Ringan-Berat)= 63**

Kriteria Inputan X6 = pH

Diketahui Atribut pH berdasarkan data training di yaitu Asam-Netral, maka :

X6 (Asam-Netral)

Angsana= 8; Mahoni= 3; Pinus= 9; Bambu= 7; Jati= 5; Trembesi=6; Jabon= 7; Johar= 5; Sengon= 8; Cempaka= 6; Bakau= 2; **Total X1 (Asam-Netral) = 66**

Kriteria Inputan X7 =Dreinase

Diketahui Atribut Drainase berdasarkan X7(Baik)

Angsana= 23; Mahoni= 25; Pinus= 23; Bambu= 27; Jati= 24; Trembesi= 22; Jabon= 19; Johar= 25; Sengon= 27; Cempaka= 15; Bakau= 0; **Total X1 (Baik) = 230**

B. Menghitung Jumlah Kasus Per Kelas  $P(X_i / C_i)$

Perhitungan Manual Jumlah dari masing-masing kelas dibagi dengan jumlah kelas yang terdapat pada data training  $P(X_i|C_i)$

$P(\text{Angsana}=\text{"jumlah kelas"})$	$= 25/249 = 0,100401606$
$P(\text{Mahoni}=\text{"jumlah kelas"})$	$= 27/249 = 0,108433735$
$P(\text{Pinus}=\text{"jumlah kelas"})$	$= 24/249 = 0,096385542$
$P(\text{Bambu}=\text{"jumlah kelas"})$	$= 30/249 = 0,120481928$
$P(\text{Jati}=\text{"jumlah kelas"})$	$= 26/249 = 0,104417671$
$P(\text{Terambesi}=\text{"jumlah kelas"})$	$= 22/249 = 0,088353414$
$P(\text{Jabon}=\text{"jumlah kelas"})$	$= 19/249 = 0,076305221$
$P(\text{Johar}=\text{"jumlah kelas"})$	$= 25/249 = 0,100401606$
$P(\text{Sengon}=\text{"jumlah kelas"})$	$= 28/249 = 0,112449799$
$P(\text{Cempaka}=\text{"jumlah kelas"})$	$= 16/249 = 0,064257028$
$P(\text{Bakau}=\text{"jumlah kelas"})$	$= 7/249 = 0,028112450$

Karena setiap atribut memiliki beberapa kriteria, maka dilakukan perhitungan untuk mengetahui nilai pada setiap kriteria aribut sebagai berikut :

a. X1 = Ketinggian (Mdpl)  $P(X1|Xtotal)$

Diketahui Atribut Ketinggian berdasarkan tabel 4.2 data training di atas yaitu 225 Mdpl, maka :

X1 (225)  $P(X1|Xtotal)$

$P(\text{Angsana} \rightarrow (0/1) = 0; P(\text{Mahoni} \rightarrow (0/1) = 0; P(\text{Pinus} \rightarrow (0/1) = 0; P(\text{Bambu} \rightarrow (1/1) = 1; P(\text{Jati} \rightarrow (0/1) = 0; P(\text{Trembesi} \rightarrow (0/1) = 0; P(\text{Jabon} \rightarrow (0/1) = 0; P(\text{Johar} \rightarrow (0/1) = 0; P(\text{Sengon} \rightarrow (0/1) = 0; P(\text{Cempaka} \rightarrow (0/1) = 0; P(\text{Bakau} \rightarrow (0/1) = 0;$

b. X2 = Curah Hujan (mm/thn)  $P(X2|Xtotal)$

Diketahui Atribut Curah hujan berdasarkan tabel 4.2 data training di atas yaitu 1070 mm/thn, maka :

X2 (1070)

$P \rightarrow (0/1) = 0; P(\text{Mahoni} \rightarrow (0/1) = 0; P(\text{Pinus} \rightarrow (0/1) = 0; P(\text{Bambu} \rightarrow (1/1) = 1; P(\text{Jati} \rightarrow (0/1) = 0; P(\text{Trembesi} \rightarrow (0/1) = 0; P(\text{Jabon} \rightarrow (0/1) = 0; P(\text{Johar} \rightarrow (0/1) = 0; P(\text{Sengon} \rightarrow (0/1) = 0; P(\text{Cempaka} \rightarrow (0/1) = 0; P(\text{Bakau} \rightarrow (0/1) = 0;$

c. X3 = Temperatur (°C)  $P(X3|Xtotal)$

Diketahui Atribut Temperatur berdasarkan tabel 4.2 data training di atas yaitu 37°C, maka :

X3 (37°)

$P(\text{Angsana} \rightarrow (0/2) = 0; P(\text{Mahoni} \rightarrow (0/2) = 0; P(\text{Pinus} \rightarrow (0/2) = 0; P(\text{Bambu} \rightarrow (1/2) = 0,5; P(\text{Jati} \rightarrow (0/2) = 0; P(\text{Trembesi} \rightarrow (1/2) = 0,5; P(\text{Jabon} \rightarrow (0/2) = 0; P(\text{Johar} \rightarrow (0/2) = 0; P(\text{Sengon} \rightarrow (0/2) = 0; P(\text{Cempaka} \rightarrow (0/2) = 0; P(\text{Bakau} \rightarrow (0/2) = 0;$

d. X4 = Kemiringan (%)  $P(X4|Xtotal)$

Diketahui Atribut Kemiringan berdasarkan tabel 4.2 data training di atas yaitu 41%, maka :

X4 (41%)

$P(\text{Angsana} \rightarrow (0/3) = 0; P(\text{Mahoni} \rightarrow (0/3) = 0; P(\text{Pinus} \rightarrow (0/3) = 0; P(\text{Bambu} \rightarrow (1/3) = 0,333333; P(\text{Jati} \rightarrow (2/3) = 0,666667; P(\text{Trembesi} \rightarrow (0/2) = 0; P(\text{Jabon} \rightarrow (0/2) = 0; P(\text{Johar} \rightarrow (0/2) = 0; P(\text{Sengon} \rightarrow (0/2) = 0; P(\text{Cempaka} \rightarrow (0/2) = 0; P(\text{Bakau} \rightarrow (0/2) = 0;$

e. X5 = Tekstur  $P(X5|Xtotal)$

Diketahui Atribut Tekstur berdasarkan tabel 4.2 data training di atas yaitu Ringan- Berat, maka :

X5 (Ringan-Berat)

$P(\text{Angsana} \rightarrow (4/63) = 0,063492; P(\text{Mahoni} \rightarrow (6/63) = 0,095238; P(\text{Pinus} \rightarrow (4/63) = 0,063492; P(\text{Bambu} \rightarrow (7/63) = 0,111111; P(\text{Jati} \rightarrow (4/63) = 0,063492; P(\text{Trembesi} \rightarrow (5/63) = 0,079365; P(\text{Jabon} \rightarrow (4/63) = 0,063492; P(\text{Johar} \rightarrow (12/63) = 0,190476; P(\text{Sengon} \rightarrow (10/63) = 0,158730; P(\text{Cempaka} \rightarrow (2/63) = 0,031746; P(\text{Bakau} \rightarrow (5/63) = 0,079365;$

f. X6 = pH  $P(X6|Xtotal)$

Diketahui Atribut pH berdasarkan tabel 4.2 data training di atas yaitu Asam-Netral, maka :

X6(Asam-Netral)

$P(\text{Angsana} \rightarrow (8/66) = 0,121212121; P(\text{Mahoni} \rightarrow (3/66) = 0,045454545; P(\text{Pinus} \rightarrow (9/66) = 0,136363636; P(\text{Bambu} \rightarrow (7/66) = 0,106060606; P(\text{Jati} \rightarrow (5/66) = 0,075757576; P(\text{Trembesi} \rightarrow (6/66) = 0,090909091; P(\text{Jabon} \rightarrow (7/66) = 0,106060606; P(\text{Johar} \rightarrow (5/66) = 0,075757576; P(\text{Sengon} \rightarrow (8/66) = 0,121212121; P(\text{Cempaka} \rightarrow (6/66) = 0,090909091; P(\text{Bakau} \rightarrow (2/66) = 0,03030303;$

g. X7 =Dreinase  $P(X7|Xtotal)$

Diketahui Atribut Dreinase berdasarkan tabel 4.2 data training di atas yaitu Baik, maka :

X7(Baik)

$P(\text{Angsana} \rightarrow (23/230) = 0,1; P(\text{Mahoni} \rightarrow (25/230) = 0,108695652; P(\text{Pinus} \rightarrow (23/230) = 0,1; P(\text{Bambu} \rightarrow (27/230) = 0,117391304; P(\text{Jati} \rightarrow (24/230) = 0,104347826; P(\text{Trembesi} \rightarrow (22/230) = 0,095652174; P(\text{Jabon} \rightarrow (19/230) = 0,082608696; P(\text{Johar} \rightarrow (25/230) = 0,108695652;$



PSengon  $\rightarrow (27/230) = 0,117391304$ ; PCempaka  $\rightarrow (15/230) = 0,065217391$ ; PBakau  $\rightarrow (0/230) = 0$ ;  
 Kalikan Semua Variabel Kelas “Pi(Xi) x P(Ci)”

Berdasarkan Data testing yang ada probabilitas dari setiap variabel, atribut dan probabilitas dikalikan untuk melihat nilai dari setiap variabel, sehingga :

PAngsana  
 $= (0 \times 0 \times 0 \times 0 \times 0.063492063 \times 0.121212121 \times 0.1) \times (0.10401606) = 0$

PMahoni  
 $= (0 \times 0 \times 0 \times 0 \times 0.95238095 \times 0.045454545 \times 0.108695652) \times (0.108433735) = 0$

PPinus  
 $= (0 \times 0 \times 0 \times 0 \times 0.063492063 \times 0.136363636 \times 0.108695652) \times (0.096385542) = 0$

PBambu  
 $= (1 \times 1 \times 0.5 \times 0.333333 \times 0.111111111 \times 0.106060606 \times 0.117391304) \times (0.120481928) = \mathbf{2.7791}$

PJati  
 $= (0 \times 0 \times 0 \times 0.666667 \times 0.063492063 \times 0.075757576 \times 0.104347826) \times (0.104417671) = 0$

PTerambesi  
 $= (0 \times 0 \times 0.5 \times 0 \times 0.079365079 \times 0.090909091 \times 0.95652174) \times (0.088353414) = 0$

PJabon  
 $= (0 \times 0 \times 0 \times 0 \times 0.063492063 \times 0.106060606 \times 0.82608696) \times (0.076305221) = 0$

PJohar  
 $= (0 \times 0 \times 0 \times 0 \times 0.190476190 \times 0.075757576 \times 0.108695652) \times (0.100401606) = 0$

PSengon  
 $= (0 \times 0 \times 0 \times 0 \times 0.158730159 \times 0.121212121 \times 0.117391304) \times (0.112449799) = 0$

PCempaka  
 $= (0 \times 0 \times 0 \times 0 \times 0.031746032 \times 0.090909091 \times 0.65217391) \times (0.064257028) = 0$

PBakau  
 $= (0 \times 0 \times 0 \times 0 \times 0.079365079 \times 0.0303030 \times 0) \times (0.028112450) = 0$

Berdasarkan hasil diatas maka diketahui hasil konsultasi yang sesuai dengan data testing yaitu Bambu karena PBambu(Xi|Xtotal) x P(Ci) memiliki nilai maksimum **2.7791** atau  $\neq 0$ , dilakukan penginputan 10 jenis data yang kemudian dilakukan

C.. Pengujian Dengan k-fold Cross-Validation

Setelah Memastikan Bahwa Proses perhitungan aplikasi dapat mengeluarkan hasil dengan Baik, Maka dilakukan tahap pengujian untuk menentukan akurasi dari Model Naive Bayes terhadap Data dan Hasil Luaran Aplikasi. Data diuji dengan 10 fold validation dengan menggunakan Aplikasi Rapid Minder.

Dari data hasil luaran sistem yang didapatkan, kemudian dilakukan pengujian untuk mencari akurasi dari model yang

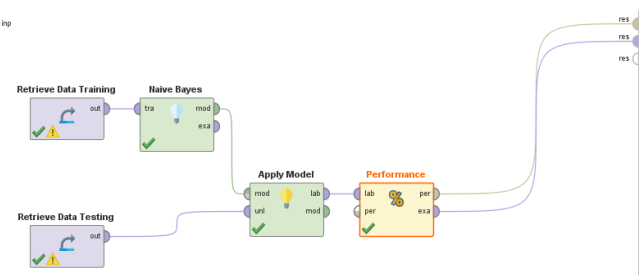
didapatkan. Pengujian dilakukan dengan 2 tahap diantaranya:

Tabel 2. Data Hasil Luaran Sistem

Jenis Tanaman	Manfaat	Ketinggian (mdpl)	Curah Hujan	Temperatur (°C)	Kelembapan	Tekstur	pH	Drainase
Angsana	Klimatologis	450	750	22	6	Ringan-Berat	Asam	Tidak Baik
Angsana	Klimatologis	680	510	31	1	Sedang	Netral	Baik
Mahoni	Klimatologis	85	4920	18	20	Ringan-Berat	Asam	Baik
Mahoni	Klimatologis	90	2000	19	34	Ringan-Berat	Basa-Netral	Tidak Baik
Pinus	Klimatologis	980	3000	17	17	Sedang	Asam-Netral	Baik
Pinus	Klimatologis	1500	2000	10	10	Ringan	Basa	Baik
Bambu	Hidrologis	970	1250	25	18	Ringan	Asam	Baik
Bambu	Hidrologis	860	2545	34	35	Ringan-Berat	Netral	Tidak Baik
Jati	Hidrologis	200	2000	27	30	Sedang	Basa	Baik
Jati	Hidrologis	300	2450	31	33	Sedang-Berat	Basa-Netral	Baik

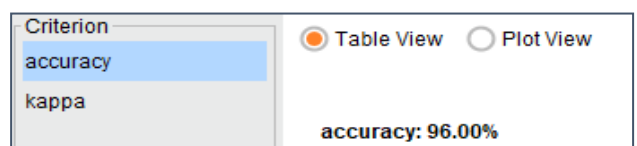
A. Pengujian Akurasi tanpa Cross Validation

Tahap pertama dilakukan dengan mengambil data hasil luaran sistem dan dijadikan data Testing sebanyak 25 data luaran. Kemudian data Training diambil dari 249 data sebelumnya.



Gambar 2. Model Pengujian Tanpa Cross Validation

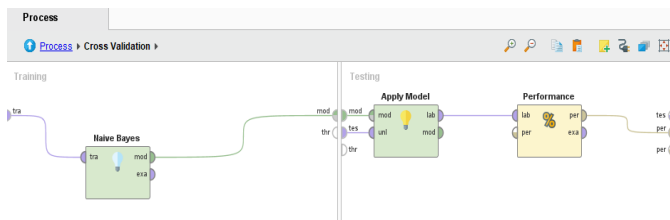
Dari Hasil Pengujian didapatkan Hasil Akurasi Sebesar 96% hal ini membuktikan bahwa akurasi dari hasil yang dikeluarkan sistem sangat Baik [] hasil luaran program yang dirancang mendapat akurasi sebesar 96%.



Gambar 3. Hasil Akurasi Tanpa Cross Validation

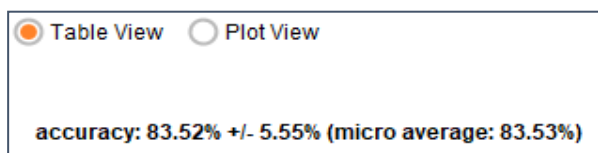
B. Pengujian dengan 10-fold Cross Validation

Pada Tahap Kedua pengujian dilakukan dengan menggunakan 10-fold Cross Validation, data hasil luaran digabung kedalam data training untuk dilakukan pengujian. Data dibagi menjadi 10 kelompok untuk nilai k pada pengujian tahap kedua.



Gambar 4. Model Pengujian 10-fold Cross Validation

Dari hasil pengujian tahap kedua didapatkan rata-rata hasil akurasi sebesar 83.53%, yang menyimpulkan bahwa model dapat mengeluarkan hasil dengan baik [16].



Gambar 5. Hasil Akurasi 10-fold Cross Validation

Berdasarkan Hasil skenario kedua tahap pengujian yang dilakukan pada penelitian ini, diperoleh tingkat akurasi skenario pertama sebesar 96% dan akurasi pengujian tahap kedua sebesar 83.53%, rata-rata akurasi dari kedua tahap pengujian didapatkan hasil sebesar 89,76%.

V. KESIMPULAN

Penerapan metode Naive Bayes dengan menggunakan parameter jenis tanaman, manfaat, ketinggian lahan, kemiringan lahan, Curah hujan lokasi, temperatur, tekstur tanah, ph tanah, dan kondisi drainase dengan 11 Class, 249 Data Training dan dilakukan pengujian dengan 2 Skenario, didapatkan hasil rata-rata akurasi sebesar 89,76% hal ini membuktikan bahwa implementasi algoritma Naive Bayes dapat digunakan untuk memilih jenis tanaman reboisasi. Untuk penelitian selanjutnya disarankan untuk menerapkan metode clustering terlebih dahulu untuk mengolah data label data training dan kemudian memilih metode lain untuk tahapan klasifikasi dan membandingkan dengan hasil luaran pada penelitian ini.

DAFTAR ACUAN

[1] Y. N. Paseneke and A. Nugroho, "Pemetaan dan Klasifikasi Kesesuaian Jenis Tanah Terhadap Tanaman Menggunakan Metode Naïve Bayes di Desa Cukilan," *J. Teknol. Inf.*, vol. 19, no. 2, pp. 199–212, 2022.

[2] J. A. Pratomo, I. S. Banuwa, and S. B. Yuwono, "Evaluasi Keberhasilan Tanaman Reboisasi pada Lahan Kompensasi Pertambangan Emas PT. Natarang Mining," *J. Sylva Lestari*, vol. 6, no. 2, p. 41, 2018.

[3] Y. A. S., "Implementasi Algoritma C4.5 Untuk Pengolahan Data Hutan Lindung (Studi Kasus: Dinas Kehutanan Sumatera Utara)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 118–122, 2019, doi: 10.30865/komik.v3i1.1577.

[4] A. Pratama, H. A. K. Aidilof, I. Saputra, M. Y. Iswand, and I., "Penerapan Model Klasifikasi Clustering Naïve Bayes Kesesuaian Lahan Tanaman," *J. Ilm. Sains, Teknol. Ekon. Sos. dan Budaya*, vol. 4, no. 4, pp. 29–34, 2020.

[5] O. K. H. Syahputra, M. Jamilah, and S. Saputra, "PROGRAM REHABILITASI HUTAN DAN LAHAN DI TAMAN HUTAN RAYA POCUT MEURAH INTAN , ACEH ( Success Rate of Enrichment Activities in Forest and Land Rehabilitation Program in Pocut Meurah Intan Forest Park , Aceh )," *J. Lingkungan. Almuslim*, vol. 1, no. 2, pp. 1–8, 2022.

[6] I. F. A. Galih Dapa Imanda, Ramadiani, "Penerapan Metode Naive Bayes Untuk Kesesuaian Lahan Tanaman Jagung," *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 5, no. 1, pp. 24–32, 2020.

[7] M. Zulfikar and H. Fahmi, "Penerapan Sistem Pendukung Keputusan Dengan Metode Naïve Bayes Dalam Menentukan Kualitas Bibit Padi Unggul Pada Balai Pertanian Pasar Miring," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 159, 2019, doi: 10.32672/jnkti.v2i2.1566.

[8] B. Eden, W. Asrul, and S. Zuhriyah, "Sistem Informasi Peramalan Harga Pangan Dengan Menggunakan Metode Naïve Bayes Di Kota Makassar," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 7–2, no. 2, pp. 163–171, 2018, doi: 10.36774/jusiti.v7i2.251.

[9] P. D. E. P. S. Winiarti, "SISTEM PENDUKUNG KEPUTUSAN MENENTUKAN JENIS TANAMAN UNTUK REBOISASI PADA DINAS KEHUTANAN DAN PERKEBUNAN KABUPATEN KEBUMEN BERBASIS WEB," *J. Sarj. Tek. Inform.*, no. Vol 5, No 2 (2017): Juni, pp. 97–106, 2017.

[10] T. Simbolon P, Syahputra T, "SISTEM PENDUKUNG KEPUTUSAN DALAM MENENTUKAN HUTAN LAYAK UNTUK DI REBOISASI PADA DINAS KEHUTANAN PROVINSI SUMATERA UTARA MENGGUNAKAN METODE WEIGHTED PRODUCT (WP)," *J. CyberTech*, vol. 1, no. 01, pp. 47–57, 2020.

[11] A. Wicaksana, "Komparasi Algoritma Klasifikasi untuk Penentuan Jenis Spesies Tanaman Hutan," *J. Cakrawala Inf.*, vol. 1, no. 1, pp. 266–276, 2021.

[12] R. T. Prasetyo and E. Ripandi, "Optimasi Klasifikasi Jenis Hutan Menggunakan Deep Learning Berbasis Optimize Selection," *J. Inform.*, vol. 6, no. 1, pp. 100–106, 2019, doi: 10.31311/ji.v6i1.5176.

[13] Meiriyama and Sudiadi, "Penerapan Algoritma Random Forest

- Untuk Klasifikasi Jenis Daun Herbal,” *Jtsi*, vol. 3, no. 1, pp. 131–138, 2022.
- [14] N. Sulistiyowati and M. Jajuli, “Integrasi Naive Bayes Dengan Teknik Sampling Smote Untuk Menangani Data Tidak Seimbang,” *Nuansa Inform.*, vol. 14, no. 1, p. 34, 2020, doi: 10.25134/nuansa.v14i1.2411.
- [15] F. Shabir, A. I. Abdullah, B. E. W. Asrul, and S. A. A. Nur, “Implementation Of The Double Exponential Smoothing Method In Determining The Planting Time In Strawberry Plantations,” *Telematika*, vol. 19, no. 2, p. 259, 2022, doi: 10.31315/telematika.v19i2.7544.
- [16] Z. Lyu et al., “Back-Propagation Neural Network Optimized by K-Fold Cross-Validation for Prediction of Torsional Strength of Reinforced Concrete Beam,” *Materials (Basel)*, vol. 15, no. 4, 2022, doi: 10.3390/ma15041477.